Banks, Jasmine (2001) A framework for a dense matching algorithm : Results using matching metrics and order statistic filters. In *Proceedings of the International Conference on Augmented, Virtual Environments and Three-Dimensional Imaging (ICAV3D'01)*, Santa Marina Hotel, Ornos, Mykonos, Greece.

A Framework for a Dense Matching Algorithm — Results Using Matching Metrics and Order Statistic Filters

Jasmine Banks

Fraunhofer Institute for Computer Graphics,
Rundeturmstr. 6, Darmstadt 64283, Germany.
Email: jbanks@ieee.org

ABSTRACT

An algorithm for computing dense correspondences between images of a stereo pair or image sequence is presented. The algorithm can make use of both standard matching metrics and the rank and census filters, two filters based on order statistics which have been applied to the image matching problem. Their advantages include robustness to radiometric distortion and amenability to hardware implementation. Results obtained using both real stereo pairs and a synthetic stereo pair with ground truth were compared.

The rank and census filters were shown to significantly improve performance in the case of radiometric distortion. In all cases, the results obtained were comparable to, if not better than, those obtained using standard matching metrics. Furthermore, the rank and census have the additional advantage that their computational overhead is less than these metrics. For all techniques tested, the difference between the results obtained for the synthetic stereo pair, and the ground truth results was small.

1. INTRODUCTION

A framework for dense matching, ie, finding correspondences for all points between two images of a stereo pair or image sequence, is presented. This is an important component of any system which produces a dense 3D reconstruction of a scene from images.

The rank and census are two filters based on order statistics which have been applied to the stereo matching problem[4]. Their advantages include their robustness to radiometric distortion and their amenability to hardware implementation. The matching algorithm presented in this paper is used to compare standard matching metrics with rank and census based techniques.

The rank and census filters are described in Section 2. Section 3 then outlines the new matching algorithm. Section 4 presents results obtained for this algorithm, using both standard matching metrics, and rank and census techniques. In addition, results obtained from matching a synthetic stereo pair are compared with a known ground truth. Section 5 provides some discussion of these results, and concludes.

This work was carried out with the support of the European Union through the CAMERA (ERB FMRX-CT97-0127) research network.

2. RANK AND CENSUS FILTERS

2.1. Rank Filter

A window of size $M \times N$ is passed over the image. At each location in the image, the number of pixels less than the centre pixel are counted. This becomes the value of the rank image at that location. Two rank filtered image regions may be compared using the SAD (Sum of Absolute Differences) metric[1].

2.2. Census Filter

Again, a window of size $M \times N$ is passed over the image. At each location in the image, the pixel neighbourhood is mapped to a bit string. If a pixel is less than the centre pixel, the corresponding position in the bit string is set to 1, otherwise it is set to 0. Regions in census filtered images may be compared by counting the number of bits which differ in the bit strings — effectively an XOR operation — and summing over the region.

3. MATCHING ALGORITHM

The overall matching algorithm is depicted in Figure 1. The various components of this algorithm are described as follows:

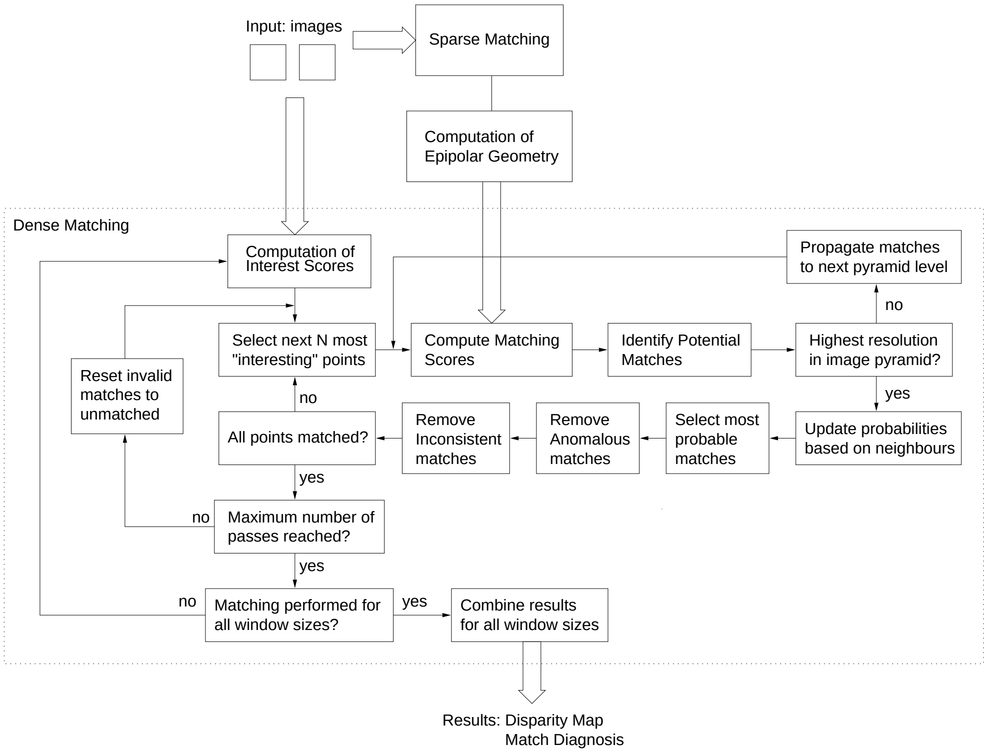


Figure 1: Overall matching algorithm.

3.1. Inputs to the Algorithm

The input to the algorithm consists of an image pair, and the fundamental matrix, which encapsulates the epipolar geometry relating the two images. The dense matching algorithm is preceded by a sparse matching stage, which automatically detects a robust set of sparse matches, and a stage which computes the epipolar geometry[5].

3.2. Computation of Interest Scores

An interest score is computed for every location in the images. Points whose interest score is below a threshold are flagged as "low interest".

3.3. Selection of Points

The N most interesting points are selected for matching, where N is typically set to one tenth the size of the image. The algorithm progressively matches the next N lesser important points, until all points are matched.

3.4. Computation of Matching Scores

A template window centered on each point in the first image is shifted in integer increments along the epipolar line in the second image. The value of the match score is computed at each candidate position, using match metrics, or the rank and census techniques. This results in an array of match scores, in which potential matches are identified as local maxima or minima, depending on the metric used. The match scores are used to estimate an initial probability for each match.

3.5. Image Pyramid

The algorithm allows for the use of an image pyramid, where match scores are computed from multiple resolution images. If the highest resolution level has not been reached, the potential matches are propagated to the next level.

3.6. Update of Match Probabilities

The match probabilities are increased in proportion to the number of neighbouring points already matched, which have a similar disparity.

3.7. Removal of Invalid Matches

For each point, the match having the highest probability is selected. The match score array is interpolated to determine the disparity to sub-pixel accuracy. A number of techniques are then used to cull matches likely to be invalid.

(a) (b)

Figure 2: Test images (a) hmmwv2 (b) corridor.

Locally anomalous matches, which differ from their neighbours more than a given threshold, are flagged as "anomalous". Matching is then performed in reverse, that is, the matched location in the second image is itself matched back to the first image. If the matched location matches back to a location other than the original point, the match is flagged as "inconsistent".

3.8. Multiple Algorithm Passes

It is possible to run the algorithm more than once — on each "pass", all points flagged as having invalid matches are set to unmatched, and it is attempted to match these points again.

3.9. Multiple Match Window Sizes

The disparity results obtained for a number of different match window sizes are combined into a single result. This aims to improve accuracy by removing the "smoothing" effect introduced by using a window of pixel values for matching.

3.10. Output of the Algorithm

The output of the algorithm consists of the x-, y- and absolute value of disparity for each point in the image that was successfully matched, as well as an image of "match flags". Each point in the "match flags" image has one of the following values:

matched - point successfully matched.

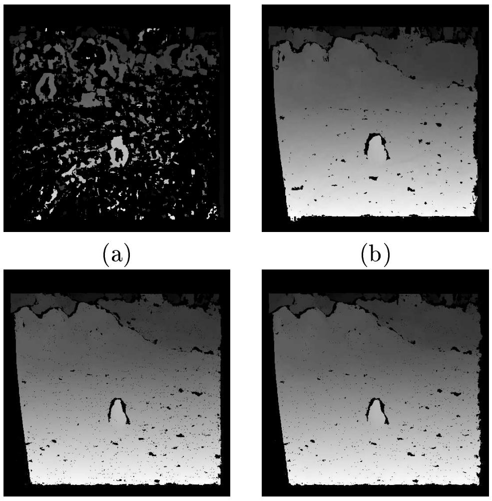
low interest - interest score below threshold.

not found - there exists no local maxima or minima of the match scores array.

inconsistent - match failed the consistency test.

anomalous - match was locally anomalous.

border - location was a border region.



(a) (b) (c) (d)

Figure 3: Disparity results for hmmwv2 stereo pair (a) SAD (b) NCC (c) RANK (d) CENSUS.

4. TESTING AND RESULTS

Figure 2(a) shows the left image of the test stereo pair "hmmwv2", obtained from [2]. This image pair suffers from radiometric distortion, the left image being approximately 15% brighter than the right.

Comparison with ground truth, ie, "correct" results, provides a method of testing the accuracy of the results returned by the matching algorithm. However, dense ground truth data is generally not available for real stereo pairs. The "corridor" stereo pair[3], for which the left image is shown in Figure 2(b), is one synthetic stereo pair for which ground truth disparity data is known.

Disparity results for the test pairs, using the SAD (Sum of Absolute Differences) and NCC (Normalised Cross Correlation) metrics [1], and the rank and census techniques, are shown in Figures 3 and 4. In each case, the results shown here were produced using the simplest case — one pyramid level, one pass of the algorithm, and only one match window size (in this case 11). For the synthetic "corridor" pair, the threshold "interest score" was set to zero, in order that matching be attempted for all image points. This was in order to focus on the comparison between standard matching metrics and rank and census based techniques, and to avoid the combinatorial explosion of alternative combinations of match algorithm input parameters.

Table 1 shows, for each test pair, the proportion of points matched, and the number of matches returned for each value of match flag.

image	match method	prop. matched	matched	low interest	not found	inconsistent	anomalous	border	rms diff.
hmmwv2	SAD	0.1849	43585	10335	22559	138294	31243	16128	-
	NCC	0.7845	184903	10335	21518	20945	8315	16128	-
	RANK	0.7915	185005	8325	21933	21104	5697	20080	-
	CENSUS	0.8010	187217	8325	22537	18310	5675	20080	-
corridor	SAD	0.8192	47182	0	1773	5795	2848	7936	0.3990
	NCC	0.7559	43539	0	3101	8044	2916	7936	0.5449
	RANK	0.8157	45431	0	4019	4796	1450	9840	0.4148
	CENSUS	0.8262	46016	0	4139	4189	1352	9840	0.4100

Table 1: Results of matching for test images.

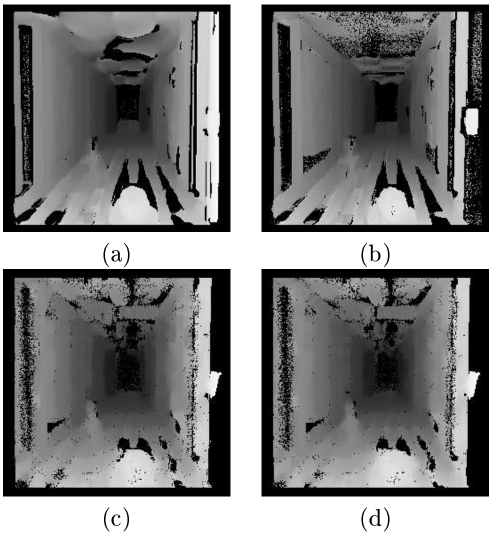


Figure 4: Disparity results for corridor images (a) SAD (b) NCC (c) RANK (d) CENSUS.

The computation of the proportion matched does not include points flagged as "border" or "low interest", for which matching was not attempted. For the "corridor" images, Table 1 also shows the computed rms difference been the disparity of the matched pixels, and the ground truth disparity.

5. DISCUSSION

From Figure 3 it can be clearly seen that the SAD is not robust to radiometric distortion. Other metrics, such as the NCC, are robust to this type of distortion, however introduce more computational overhead. It can be seen from Table 1, and from results from of other test pairs not shown in this paper, that the performance of the rank and census is generally comparable to, if not better than, metrics such as the NCC. Furthermore, they have the advantage that their computational overhead is much less than these metrics.

The matching metrics and the rank and census have also been compared using a known ground truth pair. In Table 1 shows this is the only pair for which the SAD metric has outperformed all other techniques. It is supposed that this is due to the "perfect" nature of a synthetic stereo pair — in that it does not suffer from noise or radiometric distortion. However, it can be seen from Table 1 that the rank and census have performed comparable to the SAD, and that for all matching techniques, the difference between the computed disparity and ground truth disparity is small.

Further work would involve more extensive testing of the presented matching framework — for example, using multiple "passes" of the algorithm to re-match points flagged as "not found", "inconsistent" or "anomalous", and combining results obtained from multiple window sizes.

6. REFERENCES

[1] P. Aschwanden and W. Guggenbühl. Experimental results from a comparative study on correlation-type registration algorithms. In Förstner and Ruwiedel, editors, *Robust Computer Vision*, pages 268–289. Wickmann, 1993.

[2] Vision and Autonomous Systems Center's Image Database. http://www.vasc.ri.cmu.edu/idb/.

[3] Stereo Images with Ground Truth Disparity and Occlusion. http://www-dbv.cs.uni-bonn.de/ ft/stereo.html.

[4] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proceedings ECCV'94*, pages 151–158, Stockholm, May 1994.

[5] Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. Technical Report 2927, INRIA, July 1996.